April 2009

Identification of the sit still gene – shaky suspicions about flies, fats and multiple sclerosis

Shreelekha Mandal
Worcester Polytechnic Institute

Identification of the *sit still* gene –
shaky suspicions about flies, fats, and multiple sclerosis

Author: Shreelekha Mandal

Advisor: Joseph B. Duffy

CONTENTS

ABSTRACT

Lipids are known to be important for proper cellular and metabolic functions. Lipid disorders include Tay Sach's, Gaucher's, Neimann-Pick, diabetes, adrenoleukodystrophy (ALD), and macular degeneration. Macular degeneration has been linked to mutations in a lipid elongase, of which 7 exist in humans and 20 in the fruit fly, *Drosophila melanogaster*. To understand the contributions of elongases to development, I surveyed their function in *Drosophila* using RNAi. I identified *sit still* (*sit*), whose phenotype suggests a model for neurodegenerative diseases like multiple sclerosis.

ACKNOWLEDGEMENTS

First and foremost, I owe this project to Duff. Thanks for taking me into your lab and for your guidance and patience at every step-I think it would be redundant to keep saying this- but there is no doubt that I couldn't do this without you. Thank you for all the conversations and a chance to get to know your lab, and you as an advisor and confidante. I will miss working with you.

Thank you Prachi Gupta, for helping me with absolutely everything and the incredible amount of patience you had with me. I wish you all the best for the future- you will do well in everything! Please keep in touch and take care of yourself. And the fruit flies, of course.

Everyone in lab –thanks so much for having me around. Not only was your intelligence and advice always appreciated, it was great fun to talk about all the things we did and spend time.

I would like to thank my best friends, all of them, for supporting me and believing in me always. Alice, you are my sister and thanks for making me work harder and being true to myself. Gaurav- I hope you like non-flying flies. Thanks for being my rock. Kelley- thanks for being there when I needed you. And thanks, of course, for accusing me of bringing flies into the apartment.

And of course, I would like to thank my parents for their grace, their love, and every single thing they have done for me. They are the reason I am here and that I am able to do this. I can only hope and pray that I am able to make you proud and content. I love you.

INTRODUCTION

With an expanding world population, human genetic disorders have become more prevalent in society today. Of these, neurodegenerative diseases form a major constituent. Such diseases are characterized by the progressive loss of neural tissue over time and pose a growing health care challenge to an aging population. For this reason, research on developing therapies and finding the molecular causes behind many of these diseases has increased dramatically.

Often neurodegenerative disorders are caused by lipid-related defects in the body. One such example is Tay-Sach's disease, a lipid storage disorder associated with mental and physical disabilities leading to fatality. The underlying defect is insufficient activity of beta-hexosaminidase A, the enzyme that breaks down gangliosides in the body as needed, which causes the accumulation of fatty acid derivatives called gangliosides in the body leading to fatality caused by the accumulation of fatty acid derivatives called gangliosides in the body (Okada S & O'Brien JS, 1969).

1.1 Fatty Acids and Elongases

In all eukaryotic cellular systems, fatty acids, molecules composed of a long hydrocarbon chain and a carboxylic acid group, are known to be vital for proper function (Figure 1). They may be saturated or unsaturated, and are available to organisms via their diet, or via cellular synthesis in the cytosol (Figure 1). Within the endoplasmic reticulum, cytoplasmic fatty acids with 16 carbon units (C_{16}) can be converted into long chain fatty acids (LCFAs: $C_{18\text{-}23}$) or very long chain fatty acids (VLCFAs: $C_{24\text{-}26}$), by the addition of two carbon units at a time at the carboxyl end (Figure 1). Fatty acids function in a plethora of cellular and metabolic functions. For example, the hydrophobicity of fatty acids allows for enough forces to create the lipid bilayer membrane of cells. Moreover, the breakdown of fatty acids releases great amounts of energy, primarily in the form of acetyl-CoA, which is used as a building block for lipids and proteins, or further broken down in the citric acid cycle into amino acids or fatty acids. Finally, fatty acids are often components of other compounds like hormones and signaling molecules (Tehlivets O, Scheuringer K, & Kohlwein SD, 2007).

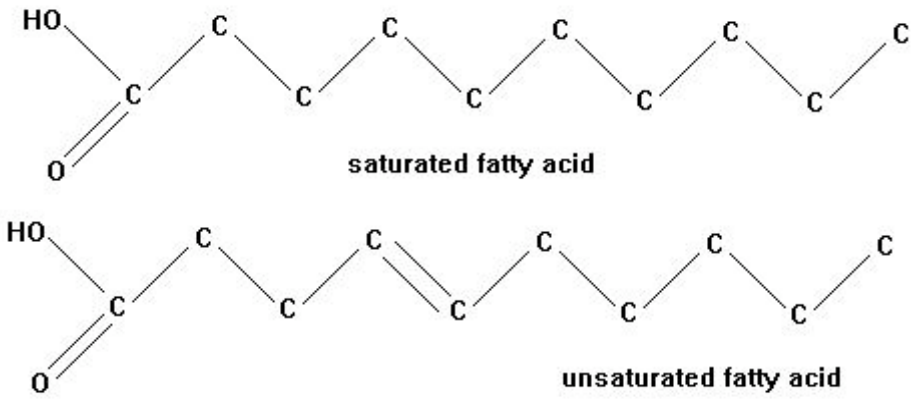

Figure 1: Saturated vs. unsaturated fatty acids. The former has only single bonds, and is 'saturated' with hydrogen. Only the latter can have additions to become a longer fatty acid

Fatty acids synthesized in the cytosol or taken up in the diet can be converted into VLCFAs, which are often esterified (joined by ester or amide linkages) to build sphingolipids, glycerophospholipids, triacylglycerols, and, sterol and wax esters. VLCFAs are also thought to be important for cellular and metabolic functions, including intracellular signaling and transport (Jakobsson A, Westerberg R, & Jacobsson A, 2006).

In mammals, fatty acid elongation is a four-step process that cycles until a specified length is achieved. The first step of this process is believed to be a regulatory condensation step, performed by enzymes termed elongases. These are encoded for by a set of genes called *Elovl* (elongation-of-very-long-chain-fatty-acids) and current evidence suggests that all elongases utilize a C_{16} substrate to carry out enzyme-specific elongation leading to distinct fatty acids. (Jakobsson A et al., 2006). At least six members of the enzyme family exist in mice and humans, while *Drosophila* has 20 family members; and at least one, *elovl4* in humans, has been linked to an inherited disorder - Stargadt-like macular degeneration (Li W et al., 2007).

As discussed, VLCFAs are commonly needed in building lipids required in signaling systems. One such example is the pathway for sphingolmyelin, a fatty acid enriched in the myelin sheath surrounding neuronal axons, which regulates cell survival (Ariga T, Jarvis WD, & Yu RK, 1998). This provides an important clue as to why several neurodegenerative diseases,

including Niemann-Pick, Gaucher's and Fabry's diseases, stem from lipid-related disorders in the body.

Six of the mammalian elongases, termed ELOVL1-6, have been identified to date. In general, elongases can elongate saturated and monosaturated, or polyunsaturated fatty acid chains. Consistent with the model that they produce distinct end-products, spatially and temporally, they exhibit specific expression patterns throughout the body (Jakobsson A et al., 2006). Table 1 below summarizes what is known about them to date:

GENE	PROTEIN	FUNCTION
elovl1	ELOVL1	• Related to the synthesis of C_{26} saturated fatty acids. • Reduction in *elovl1* expression increases elongation defects.
elovl2	ELOVL2	• Thought to play a role on the synthesis of polyunsaturated fatty acids. • Function believed to overlap with that of *elovl5*.
elovl3	ELOVL3	• Thought to elongate saturated C_{16-22} fatty acids during initial exposure to cold. • Needed in brown adipose tissue to form lipid droplets when body temperature suddenly drops. • Believed to perform a skin-barrier function to prevent desiccation. (Westerberg R et al., 2004)
elovl4	ELOVL4	• Point mutation linked to macular degeneration and Stargardt-like dystrophy. • May be involved in the synthesis of docosahexanoic acid. • Expression was also found in thymus and skin. • Knockout mutation causes perinatal lethality (Li W et al., 2007 Feb 7)
elovl5	ELOVL5	• Thought to be involved in elongation of C_{18-20} polyunsaturated fatty acids.
elovl6	ELOVL6	• Elongates C_{12-16}saturated fatty acids to C_{18} but not beyond that. • Found to be expressed ubiquitously, especially in high lipid-containing cells. • Deficiency causes insulin resistance and diabetes. (Matsuzaka T et al., 2007)

Table 1: List of human elongases and their known functions (review: Jakobsson et al., 2006)

1.2 Model Organisms- Drosophila:

Drosophila melanogaster provides as a good model organism to study human diseases and develop potential therapeutics. Greater than sixty percent of the human genes are conserved in this organism (Rubin GM, 2000). Many of the extracellular and intracellular signaling pathways in *Drosophila* are similar to those of humans (e.g. MAP kinase pathway) (Bier E, 2005). Moreover, they are small, have short life cycles and are easy to store; and both genetic and compound screens can also be performed easily on them.

In addition, a wealth of genetic tools exists to manipulate gene function, including the UAS-GAL4 system, which has been used extensively for the purposes of this project. Both the responder (UAS) strain and the driver (GAL4) strain- when crossed can trigger misexpression of a specified target gene and lead to phenotypes that can then be studied (for review see Duffy JB, 2002).

So far, relatively little is known about *elongases* in *Drosophila*, even in relation to human *elongases*. Limited studies have linked only 4 out of 20 fly *elongases* to their function via both loss and gain of function experiments (Table 2).

ELONGASE/GENE	PUTATIVE HUMAN ORTHOLOG	FUNCTION
baldspot/noa	ELOVL6/ ELOVL3	• Expressed in the embryonic nervous system, in imaginal discs, fat bodies, malphighian tubules and sex gonads • Reduced levels lead to impaired motility and severely reduced viability • Needed for male germline development (Jung A, Hollmann M, & Schäfer MA, 2007)
eloF	unknown	• RNAi knockdown triggers reduced copulation frequency • Crucial for female pheromone synthesis (Chertemps T et al., 2007)
elo68α	ELOVL1	• Expressed in male genital system • Important for synthesis of vaccenyl acetate (male pheromone) (Chertemps T, Duportets L, Labeur C, & Wicker-Thomas C, 2005)
bond	ELOVL7/ ELOVL4	• Loss of function causes failure of dividing spermatocytes to cytokinese (Szafer-Glusman E et al., 2008)

Table 2: List of fly elongases and their known functions

Even with the limited knowledge about this set of genes, it appears that *elongases* are critical for a variety of functions, ranging from skin protection in mammals to fertility and pheromone production in insects. The specific objective of this project aims to screen a subset of *elongases* in flies with closest homology to human *elongases* and test their functions in terms of fertility, viability, development and neuronal behavior.

1.3 The UAS-GAL4 system

The strategy used for testing of *elongase* function in this proposal was the UAS-GAL4 system. GAL4 is a transcription activator found in the yeast *Saccharomyces cerevisiae*. GAL4 binds to the UAS (upstream activating sequences) element, which activates transcription of the downstream gene. Thus, the UAS element, which has GAL4 binding sites, can control the expression of a gene of interest (called the responder) through its presence. When GAL4 is absent in responder lines, the gene of interest remains transcriptionally silent. This is represented in figure 2a. To activate transcription of the gene of interest, the responder lines can be mated to the driver lines, strains expressing GAL4 in a tissue specific manner (such as muscle, neuronal, wing, etc.). This is shown in figure 2b. The progeny from this cross will often express the responder in a transcriptional pattern similar to the driver, i.e. the responder will be expressed in a tissue-specific manner like the driver (for review see Duffy JB, 2002).

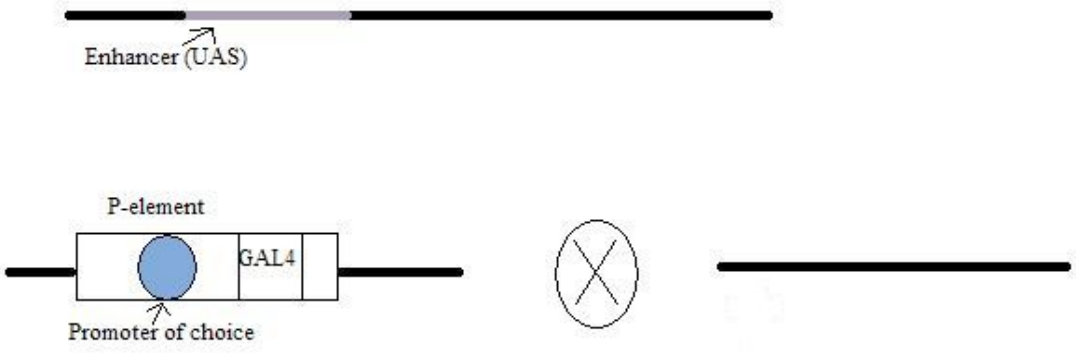

Figure 2a: No transcriptional activation, since GAL4 has no binding site in the cross

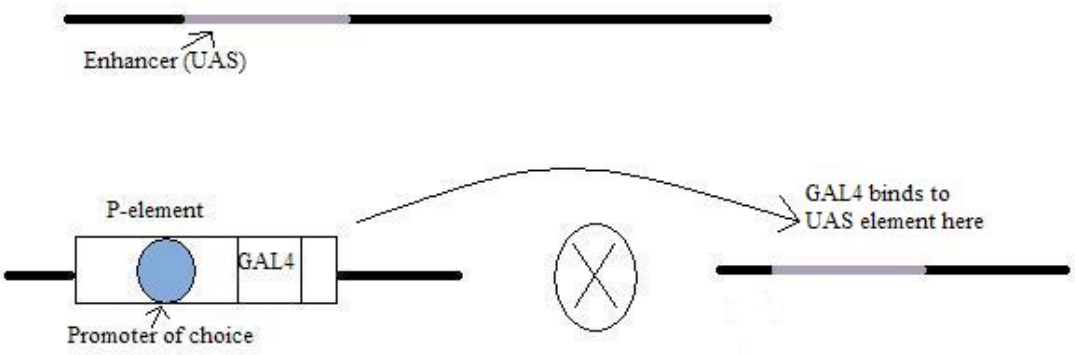

Figure 2b: Transcriptional activation of responder caused by GAL4 binding to UAS

Thus, GAL4 tissue-specific drivers can be used to express genes in specific areas of the body. The system is also temperature specific, exhibiting maximal activity at 29 °C and minimal at 16 °C (Duffy JB, 2002). For the purpose of this project, several different tissue-specific drivers were used to 'knock down' elongases using RNA-mediated gene interference, also called RNAi.

1.4 RNAi as a genetic tool

In some systems, the presence of double-stranded RNA (dsRNA) can trigger what is called RNA interference. dsRNA molecules can be processed into small interfering RNA (siRNA) by the enzyme Dicer. These siRNAs are in turn bound to a helicase and other proteins to form the RISC (RNA-induced silencing) complex. Through sequence complementarity, the siRNA guides the association of the RISC complex to target mRNA, which is then degraded by the RISC complex in a sequence specific, siRNA-dependent fashion (Pratt AJ & Macrae IJ, 2009) (Figure 3).

To assess the function of elongases in *Drosophila*, this system was combined with the UAS-GAL4 system in this project. By ordering strains expressing gene-specific dsRNAs, target elongases were 'knocked down' in a tissue specific manner by using specific GAL4 drivers. Thus, the role of specific elongases could be assessed throughout development in specific tissues of the fly body.

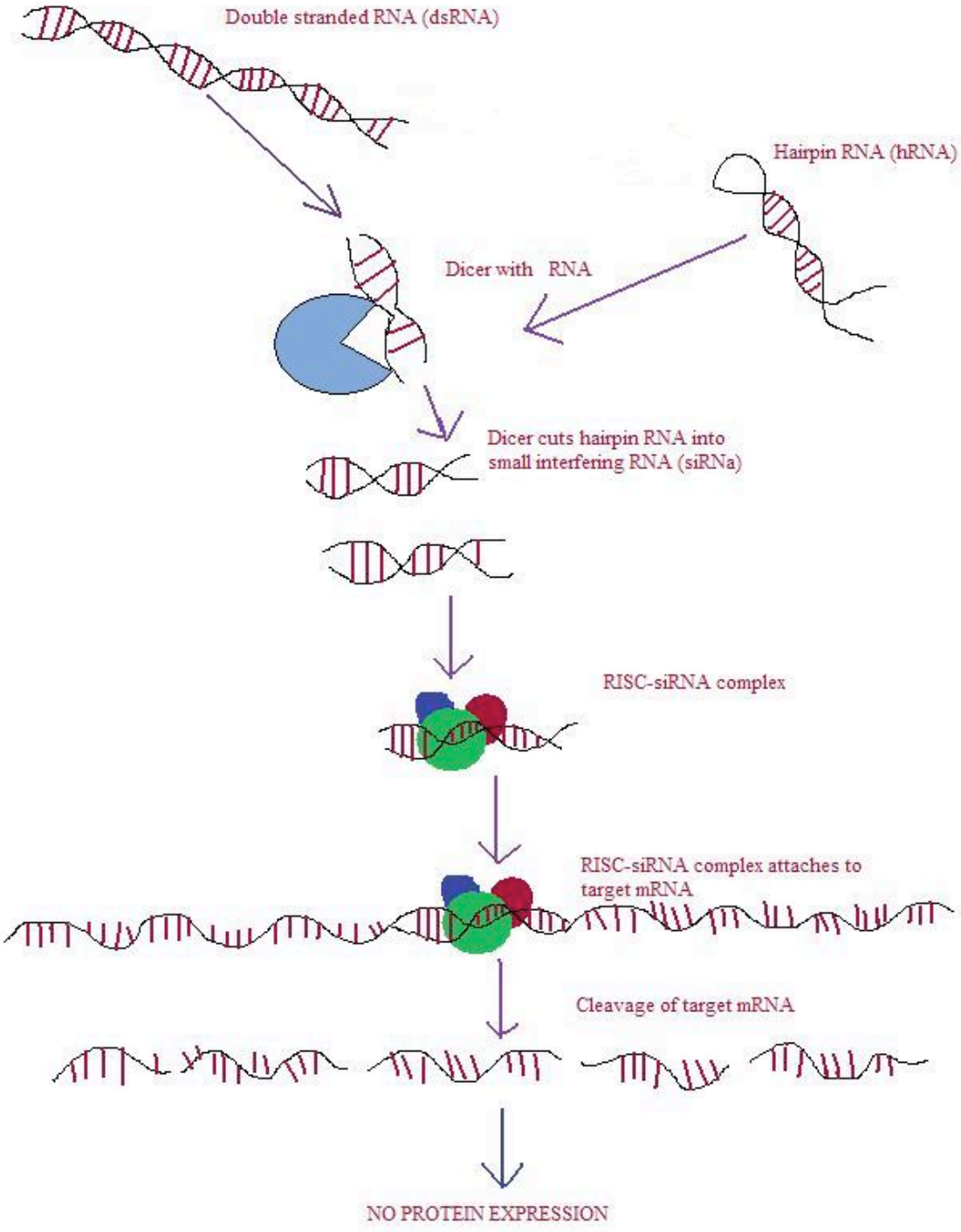

Figure 3: RNA interference. dsRNA gets broken up into siRNA by Dicer. The RISC-siRNA complex forms and degrades the target mRNA, 'knocking down' protein expression of that mRNA

MATERIALS AND METHODS

As discussed before, there are 20 elongases in the *Drosophila* proteome. Research in the Duffy lab was already being carried out on the *baldspot* (*bals*) gene, also called *noa*. Loss of function experiments in the lab demonstrated that *bals* is vital for female fertility and embryonic viability. In contrast, misexpression of *bals* led to a rough-eye phenotype, likely due to the induction of programmed cell death.

2.1 Selection of RNAi lines

Apart from the role of *bals*, I wanted to address the function and specificity of additional elongases. By performing a comparative study of elongases, I identified 5 *Drosophila* elongases sharing the greatest similarity to human elongases. Each human elongase was compared to the Drosophila proteome using the Basic Local Alignment Search Tool (BLAST) in the NCBI database. The top elongases identified in *Drosophila* were picked to carry out experiments. Strains expressing dsRNA specific to each of these elongases (RNAi lines) were then obtained ordered from the Vienna Drosophila Stock Center. The gene sequences for each of these RNAi lines are given in the Appendix.

2.2 Genetic Crosses

Six elongases, including *bals* as a control, were tested with various GAL4 drivers to test cell type specificity. These were CY2GAL4, ActGAL4, TubGAL4, C155GAL4, A9GAL4, EnGAL4, GMRGAL4, MS1096, 040GAL4, P{GawB}OK376, P{GawB}T80 and P{GawB}C179. Virgins from the GAL4 drivers were crossed to the following RNAi UAS responder lines: ELO9807 (CG31523), ELO37328 (CG31522), ELO30179 (CG6921), ELO48138 (CG2781), ELO43091 (CG5278), and *bals*47520 (CG3971). Crosses were maintained at 20ºC, 25ºC and 28ºC, to control GAL4 activity as needed. Progeny were collected and sorted by phenotype, and examined for phenotypic effects. Wild-type flies, w^{1118}, were also maintained and crossed to appropriate drivers as needed for experimental controls.

The CG31522 line was crossed to additional GAL4 drivers for further specification. These were P{GawB}477, P{GawB}D42, {GawB}109(2)80, P{GAL4-sli.S}, P{GAL4-Mhc.W}, P{GAL4-Mef2.R}, and P{GawB}how^{24B}.

2.3 Behavioral Assays

Progeny from crosses between the CG31522 RNAi line and neural GAL4 drivers were tested to check their movement and ability to orient properly. For this purpose, a set number of flies were placed into an empty vial, vortexed for 10 seconds, and observed for a further 15 seconds to see how many were able to upright themselves and climb up the sides of the vial. The progeny were kept at room temperature and tested between 2-4 days of eclosion. Progeny of complementation crosses between CG31522 mutations were also tested.

2.4 Time Shift Assays

To determine the state of development when CG31522 function is required, temperature shift experiments were carried out. For this, we had four vials of the cross between the CG31522 RNAi responder and C155GAL4: two at 20ºC, one at 25ºC and one at 28ºC. Vials were flipped everyday and kept at their respective temperatures. Each vial from 20ºC was shifted to the two higher temperatures after each successive day of development, i.e. at 24 hours, then 48 hours, then 76 hours, and so on. The vials at the higher temperatures were also shifted down to 20ºC in the same manner. This was done to increase GAL4 activity and consequently CG31522 knockdown at higher temperatures, and decrease GAL4 activity and CG31522 knockdown at 20ºC. Flies collected all vials were assessed in the behavioral assay above to characterize any behavioral defects.

2.5 Loss of Function crosses

Existing mutation and deficiency lines for CG31522 were searched for at the Bloomington Drosophila Stock Center. Two of these were transposable element inserts: P{wHy}DG02703 and P{SUPor-P}CG31522^{KG01294}, at different sites in CG31522. The other stock (BS#9224) obtained was a chromosomal deletion Df(3R)ED5071encompassing CG31522. All of these lines were then used in *inter se* complementation crosses and assessed for homozygous viability.

2.6 Generation of DNA constructs

i) Ordering cDNA

cDNAs for all the selected genes (except *bals*) were identified on FlyBase: GH22993 for CG31522, RE58951 for CG5278, GH 09808 for CG31523, LD14839 for CG6921 and RE49985 for CG2781;and ordered from the Drosophila Genomic Resource Center. Clones were transformed, grown on appropriate antibiotic plates, miniprepped and checked via restriction enzyme analysis on agarose gels, to ensure the correct clones were received. GH 22993 and GH 09808 were checked on chloramphenicol plates and all others on ampicillin plates. Flies for all cDNA clones were generated in Gene Construction Kit.

ii) Generating the Clone

To create a responder line for CG31522, Gateway cloning was carried out. For this, the CG31522 template GH22993 was used (see Appendix). To create an entry clone using site-specific recombination, primers (W100 and W101) to the coding region were designed in GCK. PCR was then used to amplify CG31522 with attB1 and attB2 recombination sites on the 5' (W100) and 3' (W101) ends respectively. With the help of Prachi Gupta, a modified PCR product with the CG31522 gene and attB1.1 and B2.1 sites was also created.

Thus, as shown in figure 4, the first step in Gateway cloning was to generate the PCR product. The next step involved using a donor vector, a 'P donor' to undergo a BP reaction with the gel purified PCR product to form the entry clone. The P donor has attP1 and P2 sites that recombine with the attB1 and B2 sites (or B1.1 and B2.1), forming the entry clone, which contained the gene of interest (CG31522) flanked by attL1 and L2 sites. Finally, an LR reaction was carried out where the entry clone was recombined with the destination vector (pUASTa6Xhis/V5 with attR1 and R2 sites). Recombination in this case gave the final expression clone - CG31522 under the control of UAS sequences and flanked by attB1 and B2 sites.

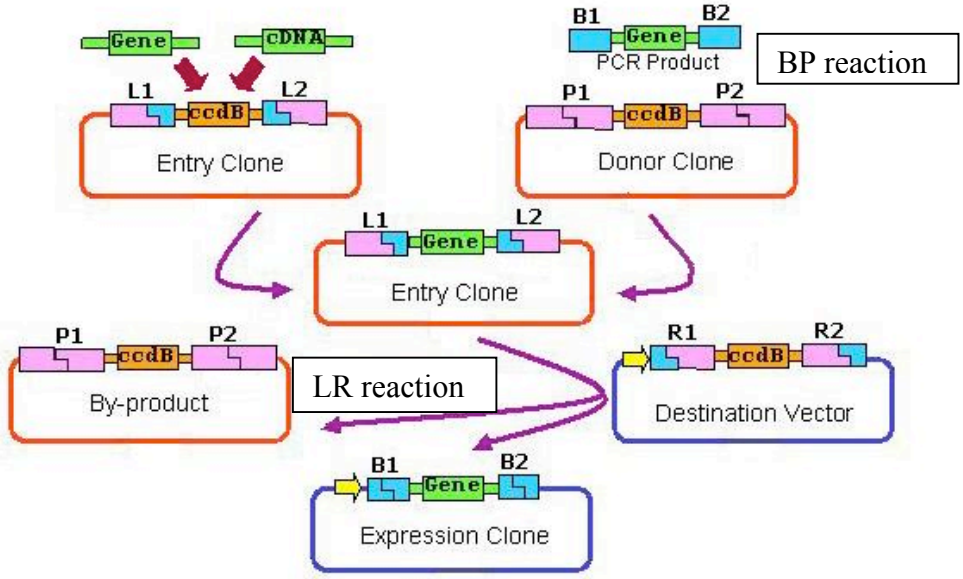

Figure 4: Gateway cloning and its steps

At every step of the entire process, every reaction was followed by transformation on plates, then minipreps of picked colonies and restriction digestion analysis. Qiagen kits were used for all mini-preps, maxi-preps, PCR purification and gel extraction as needed. After confirming the sequence of the final UAS-CG31522 clone, this DNA was sent to Genetic Services, Inc. for injection into *Drosophila* embryos and generation of transgenic flies.

2.7 Generation of UAS-CG31522 transgenics

After the DNA had been maxiprepped and sent to Genetic Services, Inc. for injection into w^{1118} *Drosophila* embryos, injected larvae were sent back so that they could be screened transgenics could be performed. Two vials of injected embryos were sent back- A and B. Adults from the injected embryos (G_0) were single-pair mated to either w^{1118} females or males (3 or 4 of each at one time) and the progeny (F1) were screened for putative transgenics based on the presence of eye color. Transgenics from G_0 were assumed to be independent insertions. In the F_1 generation, progeny with eye color were picked out and separated into vials. Vial A produced 4 transgenic males and 4 transgenic females in the F1 generation, while Vial B produced 1 transgenic male and 2 females in the F1.

Extra duplicate transgenic flies obtained from the F_1 were also separated and crossed to C155mCD8::GFP females and males to initially check for any phenotype associated with

misexpression. Based on these preliminary tests high levels of CG31522 misexpression appears to be lethal.

Transgenic flies from independent insertions at the F_2 generation were also crossed to second and third chromosome balancers- *Sp/CyO* and *Ly/TM3,Sb*, respectively. Heterozygous progeny with a copy each of the transgene and the balancer (*CyO* or *TM3,Sb*) were picked and outcrossed to w^{1118} flies to map the location of the transgene. Based on this mapping, appropriately balanced stocks will then be generated.

RESULTS

3.1 Selection of RNAi lines

RNAi responder lines for the five elongases selected were obtained. Table 3 details these RNAi lines with respect to the location (chromosome) of the inserts and their viability.

Transformant ID	Gene	Chromosome	Viability
ELO 9807	CG31523	1	Viable
ELO 37329	CG31522	3	Viable
ELO 30179	CG6921	2	Viable
ELO 43091	CG5278	2	Viable
ELO 48138	CG2781	3	Viable
BALS 47520	CG3971	3	Lethal

Table 3: RNAi responder lines

3.2 Preliminary Crosses with all RNAi lines

The RNAi lines, as mentioned before, were crossed to a variety of GAL4 drivers to assess the effects of gene specific knockdown for the selected elongases. The results from those crosses are presented in Table 4.

	CG 31522	CG 31523	CG 6921	CG 2781	CG 5278	CG 3971
Cy2GAL4 (follicle cells)	Low viability Low fertility	Collapsed eggs	WT	WT	WT	WT
ActGAL4/CyO Ubiquitous	Lethality	Lethality	Lethality	Lethality	Lethality	Lethality
TubGAL4/TM3Serr (ActGFP) Ubiquitous	Lethality	Lethality	Low viability	Lethality	Lethality	Viable

C155GAL4 **Neuronal**	Abnormal behavior (sit still phenotype)	WT	WT	WT	WT	WT
A9GAL4 **Wings**	Viable	Viable	Viable	Viable	Viable	Viable
EnGAL4 **Wings**	Viable	Lethal	Viable	Viable	Viable	Viable
GMRGAL4 **Eyes**	Rough eyes	Rough eyes	Rough eyes	Rough eyes	Rough eyes	WT
MS1096 **Wings**	Viable, inward curl	Pharate lethality, extreme wing damage in escapers	Viable, little inward curl	Viable, little inward curl	Viable, slight curl	Viable, WT
040//CyOGAL4 **Eyes**	Rough eyes	Rough eyes	Rough eyes	Rough eyes	Rough eyes	WT
P{GawB}OK376 **Oenocytes**	sit still phenotype	Viable, WT	Viable, WT	Larval lethality	Viable, WT	Viable, WT
P{GawB}T80 **3rd instar imaginal discs**	Lethal	Lethal	Low viability	Larval lethal	Viable, WT	Viable, WT
P{GawB} C179 **Embryonic mesoderm, larval muscles and wing imaginal discs**	WT	WT	Pharate lethality	Larval lethal	Lethal	WT

Table 4: Preliminary Data from RNAi crosses with GAL4 drivers

3.3 Crossing CG31522 to neuronal, glial and muscle drivers

Knockdown of CG31522 was found produce interesting behavioral phenotypes with two drivers (C155 and OK376). In both cases adults underwent premature death (~3 days after eclosion) and exhibited numerous abnormal behaviors – including remaining still (no flying, minimal crawling), twitching, appearing disoriented with little response to external stimuli, and lack of feeding. To further address the role of CG31522 knockdown was generated with a number of other neuronal, glial and muscle drivers to observe any possible phenotypes. The data from those crosses is presented in table 5.

	Tissue expressed in	Observed Phenotype
ActGAL4	Ubiquitous	Lethality
ActGAL4TubGAL80	Ubiquitous	Lethality
C155GAL4	Nervous system	"sit still" phenotype
OK376	Oenocytes	"sit still" phenotype
P{GawB}477, P{UAS-mcD8::GFP.L} LL5	Subset of dendritic arborization neurons	WT
P{GawB}109(2)80, P{UAS-mcD8::GFP.L} LL5	Multiple dendritic neurons, oenocytes and chordotonal organs	WT
P{GawB} D42	Motor neurons	WT
P{GAL4-sli.S}3	Embryonic midline glial cells and MP1 neurons	"sit still" phenotype
P{GAL4}repo	Pan-glial cells, except midline glia	WT
P{GAL4-sim.S}	Ventral midline glial cells	WT
P{GawB}4G	Neuronal subset, mostly in larval ventral ganglion region	WT
P{GawB}(3)[31-1]	Embryonic peripheral nervous system, CNS, neuroblasts	WT
P{GawB}how^{24B}	Mesoderm, somatic muscles in larvae	WT
P{GAL4-Mef2.R}	Myoblasts, somatic muscles, mesoderm	WT
P{GAL4-Mhc.W}	Larval muscle system	WT

Table 5: Data from crossing CG31522 to neuronal, glial and muscle cell drivers

From the data, the behavioral phenotype presents itself only when CG31522 is knocked down in specific tissues - namely neural cells. Moreover, when the elongase is knocked down ubiquitously, lethality is observed suggesting it is also required for viability. Based on the behavioral phenotype, I have decided to tentatively name this gene *sit still* (*sits*).

3.4 Behavioral Assays

As mentioned before, the selection of CG31522 for further testing was based upon the behavioral phenotype observed upon knockdown with neuronal and midline glial drivers. The progeny from these crosses were observed to stay quite still, not flying, grooming in their position, occasionally twitching, and dying within a few days of eclosion due to the lack of eating or mating. In comparison, wild type flies and control lines (RNAi responder or GAL4 driver only) flies were very active, flying around and eating/mating normally. A behavioral assay was developed to quantify the observed phenotypic effects. Ten flies were put into an empty vial, vortexed for 10 seconds and observed for a further 15 seconds. In this way, the number of flies that uprighted themselves and climbed up could be quantified (Figure 5).

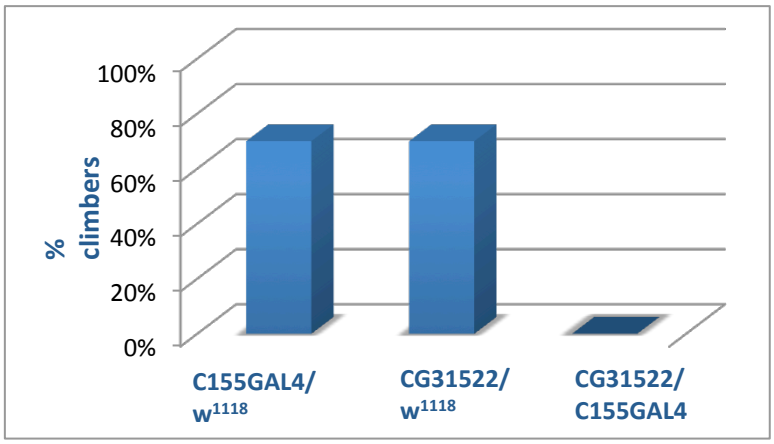

Figure 5: Behavioral effects of CG31522 knockdown (n=50)

From Figure 5, it can be seen that neural specific knockdown of CG31522 results in a dramatic effect of adult activity.

3.5 Time Shift Assays

Flies with the genotype CG31522/C155GAL4 were made to undergo the time shift assay as aforementioned (see Materials and Methods for details). Once flies that had undergone distinct temperature shifts eclosed, the behavioral assay was performed to quantify the percentage of flies

that were able to climb up the sides of the vial after the vortex experiment. The results are presented in Figure 6, a graph of an upshift from 20°C to 25°C, and a downshift of 25°C to 20°C, at every stage of development in the *Drosophila*. From this it was determined that CG31522 activity in neural cells at the pupal stage is essential for wild type adult activity.

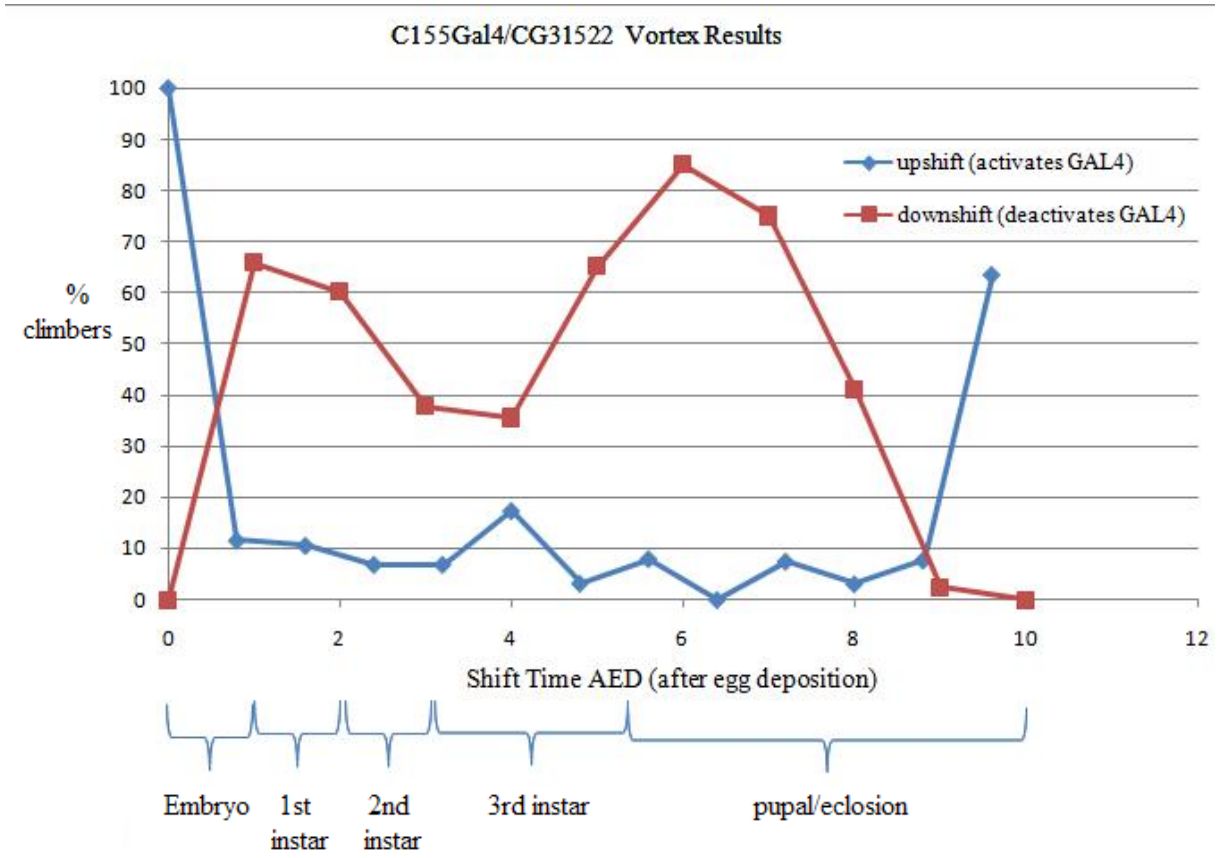

Figure 6: Developmental characteristics of CG31522 activity though temperature shifts analysis

3.6 Loss of Function crosses

Given the phenotype observed for RNAi mediated knockdown of CG31522, I sought to identify mutations in this elongase. Two transposon insertions and a large deficiency including CG31522 were identified at the Bloomington Drosophila Stock Center. The location of each transposon insertion and the deficiency is shown relative to CG31522 in Figure 7. Initially the viability of each transposon line was determined and then a series of complementation crosses was carried out to determine the relative effects of each insertion on CG31522 activity. The data

from these crosses is presented in Table 6. Of the two insertions, KG01294 is homozygous lethal, while DG02703 is homozygous viable and appeared to have no significant behavioral effects in adults. This suggested that KG01294 disrupts the activity of CG31522 and that CG31522 activity is essential for viability. Complementation tests between the transposon alleles and the deficiency resulted in lethality in both cases, indicating that DG02703 is also an allele of CG31522, albeit weaker than KG01294. It remains to be determined if the phenotype of the two alleles in combination is lethality or if this genotype would survive to adulthood and display a behavioral phenotype similar to that observed in the knockdown.

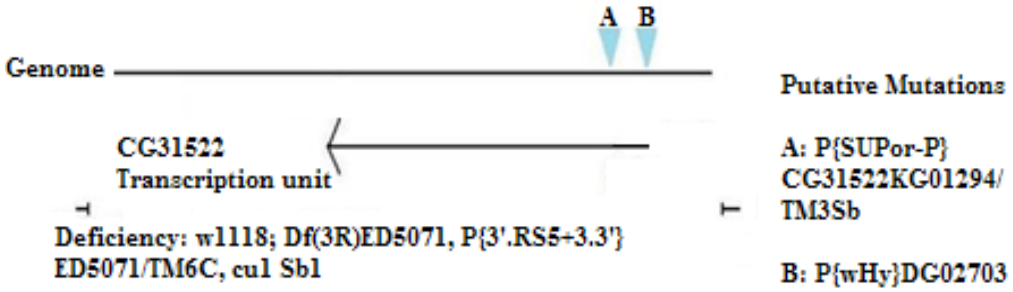

Figure 7: Genomic Region for CG31522

	P{SUPor-P} $CG31522^{KG01294}$/ TM3Sb	P{wHy}DG02703	w^{1118}; Df(3R)ED5071, P{3'.RS5+3.3'}ED5071/TM6C, cu^1 Sb^1
P{SUPor-P} $CG31522^{KG01294}$/ TM3Sb	Homozygous lethal	Not done	Lethal
P{wHy}DG02703	Not done	Viable	Lethal

Table 6: Complementation crosses for CG31522

3.6 Gain of Function Experiments

To complement the loss of function analyses presented above transgenic responder lines for gain of function studies with CG31522 were generated. From a CG31522 cDNA, a UAS-CG31522 construct was generated (Figure 8) and a total of eleven transgenic lines were made and are currently undergoing mapping and stock construction. Preliminary results with unmapped UAS-CG31522 responder lines and neural drivers that would misexpress the gene in nerves and glial cells resulted in larval lethality.

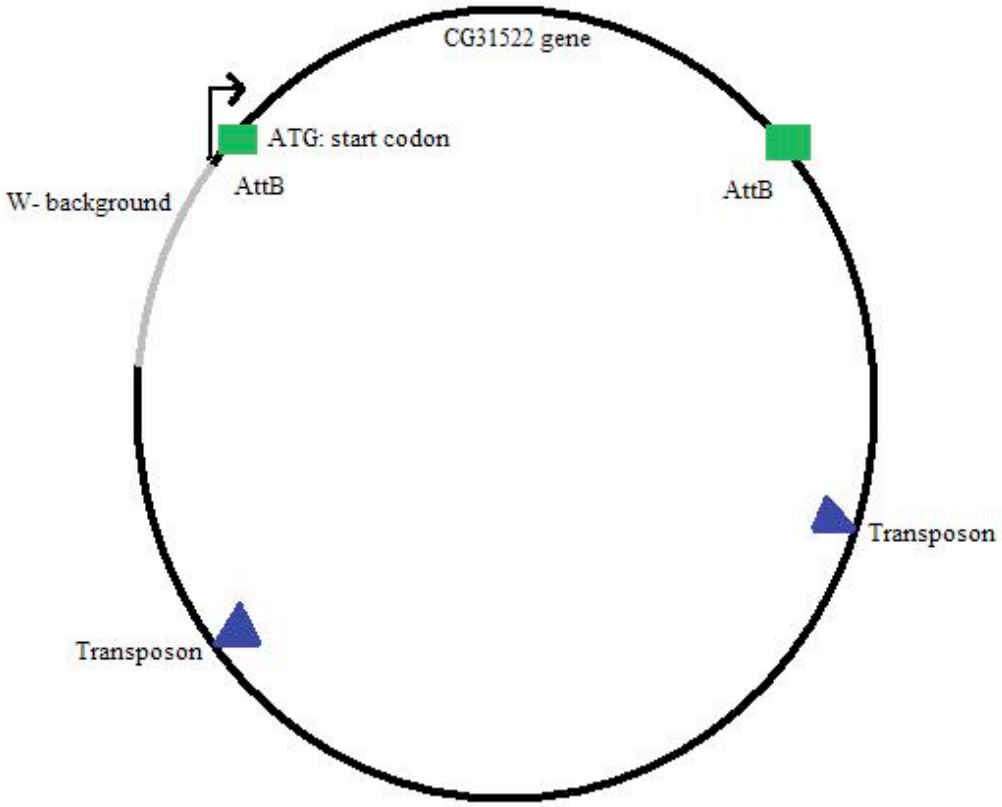

Figure 8: UAS-CG31522

DISCUSSION

4.1 Lipid-Related Disorders

Lipids have been known to be important biologically for construction of cellular membranes, intracellular trafficking, cell signaling and hormone production. However, recent studies have found links between lipids and neurodegenerative disorders such as Tay-Sachs, Niemann-Pick and adrenoleukodystrophy. In such cases, the accumulation of certain lipids in the body has been related to these diseases. On a similar note, defects in the lipid biosynthetic machinery have been linked to diseases as well, including Stargadt-like macular degeneration. Thus, the study of lipids and the biosynthetic enzymes that produce them is crucial to develop treatments or cures for such disorders.

4.2 Lipid Elongases and Development

Elongases, known to be enzymes that make fatty acid chains longer, are specific in their functions with respect to the fatty acids they produce. In relation to lipid biosynthesis, not much is known about these enzymes, apart from the fact that C_{16} fatty acid chains are made in the cytoplasm and serve as the precursor for other longer fatty acids. Elongases themselves act in the endoplasmic reticulum of cells, and are thought to play a role in the condensation step of fatty acid elongation (Jakobsson A et al., 2006). Although we have gained recent insight into the mechanism of fatty acid elongation by elongases, their roles in development still remain poorly understood. Only three of the known 7 human elongases are known to have specific functions. *ELOVL3* is known to be vital for hair and skin development, with deletion of the gene resulting in hair and skin permeability defects (Westerberg R et al., 2004). *ELOVL4* has been linked to both skin permeability barrier function - null mutations result in neonatal lethality and Stargadt-like macular dystrophy linked to a dominant ELOVL4 mutation (Li W et al., 2007). Finally, *ELOVL6* has been related to insulin resistance in obesity-induced diabetes (Matsuzaka T et al., 2007). Thus for the elongases that have been investigated, important developmental and physiological roles have been ascribed to them. It is clear that these enzymes have specific functions, but how and why this specificity exists remains to be determined.

4.3 Sit Still- fatty acids, neural development and behavior

Through my work on *sits*, it is clear that the gene is required for proper neural development. Specifically, the temperature shifts and knockdown with *sliGAL4* support the notion that *sits* is required during pupal development and in midline glial cells. This is consistent with studies that show extensive glial re-wiring takes place during metamorphosis, i.e. pupal stages.

The current phenotype of *sits* knockdown may present a working model for multiple sclerosis (MS). In MS, the damage occurs in the fatty myelin layer (produced by Schwann cells, a type of glial cell) ensheathing the axons, which may result in failure of signal propagation. Several of the symptoms are similar between MS and the *sits* phenotype - such as twitching, severe inactivity, paralysis, disorientation and premature death. Moreover, it appears that this is a developmental defect, since knockdown of *sits* in adults has no effect on behavior, whereas knockdown during development results in the classic *sits* behavioral phenotype. Since knockdown during re-wiring of the nervous system, especially that of glial cells, is crucial to produce the phenotype; one simple hypothesis is that glial cell function is compromised leading to neuronal degeneration and therefore this may be an excellent *in vivo* model for MS. At the moment, not enough evidence about the disease itself or the cellular basis for the *sits* phenotype exists to argue this conclusively. There are current studies relating MS to viral infections, inflammatory lesions and axonal loss. Arguments exist that claim both environmental and genetic factors may be responsible for the disease. However, much still remains to be uncovered (Kalman B & Lublin FD, 1999). Thus, the next function becomes to test whether the *sits* phenotype is, in fact, a good model for multiple sclerosis. Important tools to do so would be identifying loss-of-function mutations in *sits* that phenocopy the RNAi effects and gain-of-function tools for *sits* as well. I have recently achieved both of these aims and thus the lab is well positioned to test the relevance of *sits* to MS in the near future.

REFERENCES

Ariga T, Jarvis WD, & Yu RK. (1998). Role of sphingolipid-mediated cell death in neurodegenerative diseases. *Journal of Lipid Research, 39*(1), 1-16.

Bier E. (2005). Drosophila, the golden bug, emerges as a tool for human genetics. *Nature Reviews.Genetics, 6*(1), 9-23.

Chertemps T, Duportets L, Labeur C, Ueda R, Takahashi K, Saigo K, et al. (2007). A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America, 104*(11), 4273-4278.

Chertemps T, Duportets L, Labeur C, & Wicker-Thomas C. (2005). A new elongase selectively expressed in drosophila male reproductive system. *Biochemical and Biophysical Research Communications, 333*(4), 1066-1072.

Duffy JB. (2002). GAL4 system in drosophila: A fly geneticist's swiss army knife. *Genesis (New York, N.Y.: 2000), 34*(1-2), 1-15.

Jakobsson A, Westerberg R, & Jacobsson A. (2006). Fatty acid elongases in mammals: Their regulation and roles in metabolism. *Progress in Lipid Research, 45*(3), 237-249.

Jung A, Hollmann M, & Schäfer MA. (2007). The fatty acid elongase NOA is necessary for viability and has a somatic role in drosophila sperm development. *Journal of Cell Science, 120*(Pt 16), 2924-2934.

Kalman B, & Lublin FD. (1999). The genetics of multiple sclerosis. A review. *Biomedicine & Pharmacotherapy, 53*(8), 358-370.

Li W, Sandhoff R, Kono M, Zerfas P, Hoffmann V, Ding BC, et al. (2007). Depletion of ceramides with very long chain fatty acids causes defective skin permeability barrier function, and neonatal lethality in ELOVL4 deficient mice. *International Journal of Biological Sciences, 3*(2), 120-128.

Matsuzaka T, Shimano H, Yahagi N, Kato T, Atsumi A, Yamamoto T, et al. (2007). Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance. *Nature Medicine, 13*(10), 1193-1202.

Okada S, & O'Brien JS. (1969). Tay-sachs disease: Generalized absence of a beta-D-N-acetylhexosaminidase component. *Science, 165*(894), 698-700. Retrieved from PubMed database.

Pratt AJ, & Macrae IJ. (2009). The RNA induced silencing complex: A versatile gene-silencing machine. *The Journal of Biological Chemistry,*

Rubin GM. (2000). Comparative genomics of the eukaryotes. *Science, 287*(5461), 2204-2215.

Szafer-Glusman E, Giansanti MG, Nishihama R, Bolival B, Pringle J, Gatti M, et al. (2008). A role for very-long-chain fatty acids in furrow ingression during cytokinesis in drosophila spermatocytes. *Current Biology : CB, 18*(18), 1426-1431.

Tehlivets O, Scheuringer K, & Kohlwein SD. (2007). Fatty acid synthesis and elongation in yeast. *Biochimica Et Biophysica Acta, 1771*(3), 255-270.

Westerberg R, Tvrdik P, Undén AB, Månsson JE, Norlén L, Jakobsson A, et al. (2004). Role for ELOVL3 and fatty acid chain length in development of hair and skin function. *The Journal of Biological Chemistry, 279*(7), 5621-5629.

APPENDIX

CG6921

RNAi trigger

CAGTTCGTCCTCATCCTGGGCTACATGCTGACCGTGGGCGCCAAGGGCTGCAACATG
CCCAAGACACTGACCTTCTTCTTCGTGGGCAACACGGTCATCTTCCTTTACCTGTTCG
GCAATTTCTATCGCAAGACCTACAAGAAGGCCAAGTCCGTGGATGGAGGCAGTCGC
ACCACCGGCAGTAGTTTGGCTCAATCCGCTCTGCGTGCCGCCGGCGGTATGGGTTGC
ATGCCGCAGACCATGAACGCCGGTAAGCATCTGCTCCAGAATGGCCAGGTCGGAAA
GGCGTACATCGATCTGAACAACAACAGCGTGAAGCCCATGAAGCTGGAATGA

Coding sequence:

ATGACGAACTACATAAAGATTGTCGAGGAGCGAATCTCCGGCCTCTCAAAAGGAGT
GGATGAGACGGTGGACAGCTGGTTCCTAATGAGCTCACCTATGCCCGTGGTGGCCGT
GGTTCTCGTCTATCTGGCGTTTGTTCTGAAAATCGGACCCGAGTACATGAAGAATCG
CAAGCCCATGGACCTCAAGCGCATCATGGTGTTCTATAATGCCTTTCAGGTCTTGTA
CTCCATTTGGATGTGCCGCACGTCCATCCAAGAGAGCAATGTGATGGCGTCGATCTT
CTCGAAGAAGTGCGAGATTAACCGCACCAGGGAGCAGAATCTGACCCTCTATTCCG
GTGCCTGGTTCTACTTTTTCTCGAAGATCATCGATCTGCTGGACACGACGTTCTTTGT
GCTGCGCAAGAAGGACAACCAGGTGTCCTTCCTACACGTGTACCATCACACCATCAC
GGTGCTCTTCTCCTGGGGCTACCTGAAGTACGCTCCCGGCGAGCAGGGCGTGATCAT
TGGCATTCTCAACTCGGGCGTGCACATTATCATGTACTTCTACTACATGGTGGCCGC
AATGGGACCCCAGTACCAGAAGTACCTGTGGTGGAAGAAGTACATGACCAGCATCC
AGCTGATCCAGTTCGTCCTCATCCTGGGCTACATGCTGACCGTGGGCGCCAAGGGCT
GCAACATGCCCAAGACACTGACCTTCTTCTTCGTGGGCAACACGGTCATCTTCCTTT
ACCTGTTCGGCAATTTCTATCGCAAGACCTACAAGAAGGCCAAGTCCGTGGATGGA
GGCAGTCGCACCACCGGCAGTAGTTTGGCTCAATCCGCTCTGCGTGCCGCCGGCGGT
ATGGGTTGCATGCCGCAGACCATGAACGCCGGTAAGCATCTGCTCCAGAATGGCCA
GGTCGGAAAGGCGTACATCGATCTGAACAACAACAGCGTGAAGCCCATGAAGCTGG
AATGA

Protein sequence:

M T N Y I K I V E E R I S G L S K G V D E T V D S W F L M S S P M P
V V A V V L V Y L A F V L K I G P E Y M K N R K P M D L K R I M V
F Y N A F Q V L Y S I W M C R T S I Q E S N V M A S I F S K K C E I
N R T R E Q N L T L Y S G A W F Y F F S K I I D L L D T T F F V L R
K K D N Q V S F L H V Y H H T I T V L F S W G Y L K Y A P G E Q G
V I I G I L N S G V H I I M Y F Y Y M V A A M G P Q Y Q K Y L W W
K K Y M T S I Q L I Q F V L I L G Y M L T V G A K G C N M P K T L T

F F F V G N T V I F L Y L F G N F Y R K T Y K K A K S V D G G S R T
T G S S L A Q S A L R A A G G M G C M P Q T M N A G K H L L Q N G
Q V G K A Y I D L N N N S V K P M K L E

CG31522

RNAi trigger

AGAACGGAGCCCTGGCCAACGGACACGCCAAGCCCAATGGCTACTGCAAGAGCATC
AACGCCCATGACGATCTGGCCATGCCACAGACCACGGAGGCGACAGCGACAGCGAC
GCCCGCATCCAAGGCCAACGGGAGCAGCACCCCGCCGAGCAATGGACATGCCAATG
GGGTGGAGAACGTCTACAAGCAGGTGGCCAACGGCAGCGCGCACAAGGGGTCCAA
CGGAGGACTGAGCAACGGGTATGCCACCAAGCTGCTGGACGACGCCTCCCAGGAGC
TGAAACAGCGGAAGACGCCGA

Coding sequence:

ATGGCCTTAATTATGAAATACATCGACAGCATAAGCAGATACATGGACTCACATAGT
GACTCGAGGACAAAGGGCTGGCCCATGATGTCGTCCCCATTCCCCACACTGGCCGTT
TGCCTCACATATGTCTATCTGGTCAAGGTGCTGGGACCCCGATTAATGGAGAACCGA
AAGCCTCTGAATCTGCAAAACACTCTAGTCATGTACAATGCCATACAGGTTGTATTT
AGCGCATGGCTGTTTTATGAGTGCTTGATGGGCGGCTGGTGGGGCTCGTACAGCTTC
CGGTGCCAGCCGGTGGACTACACAGATAGTCCCACATCCCGGAGAATGGTTCATGC
CTGCTGGTGGTACTACTTCTCGAAGTTTACAGAATTCATGGATACGATTTTCTTCGTG
TTGCGCAAGAAGTCGAGCCAAGTTACCACACTGCACGTCATCCATCACGGCTGCATG
CCCATGTCGGTCTGGTTCGGCGTCAAGTTCACACCGGGTGGTCACAGCACCTTCTTC
GGCCTGCTCAACACCTTCGTCCACATCGTGATGTACACCTACTACATGTTCTCCGCC
ATGGGTCCGCAGTACCAGAAATACCTCTGGTGGAAGAAGTACCTGACCACTCTGCA
GATGGTCCAGTTCATCTTGATCATGGTGCACGCCTTCCAGCTGCTCTTCATCGATTGC
AACTACCCGAAGGCGTTCGTCTGGTGGATTGGCATGCATGCCGTCATGTTCTTCTTCC
TGTTCAACGAGTTCTACAAGGCAGCCTACAGGAGCCGCATGATGAAAAAGAACGGA
GCCCTGGCCAACGGACACGCCAAGCCCAATGGCTACTGCAAGAGCATCAACGCCCA
TGACGATCTGGCCATGCCACAGACCACGGAGGCGACAGCGACAGCGACGCCCGCAT
CCAAGGCCAACGGGAGCAGCACCCCGCCGAGCAATGGACATGCCAATGGGGTGGA
GAACGTCTACAAGCAGGTGGCCAACGGCAGCGCGCACAAGGGGTCCAACGGAGGA
CTGAGCAACGGGTATGCCACCAAGCTGCTGGACGACGCCTCCCAGGAGCTGAAACA
GCGGAAGACGCCGAAATAA

Protein Sequence:

M A L I M K Y I D S I S R Y M D S H S D S R T K G W P M M S S P F P
T L A V C L T Y V Y L V K V L G P R L M E N R K P L N L Q N T L V
M Y N A I Q V V F S A W L F Y E C L M G G W W G S Y S F R C Q P V
D Y T D S P T S R R M V H A C W W Y Y F S K F T E F M D T I F F V
L R K K S S Q V T T L H V I H H G C M P M S V W F G V K F T P G G
H S T F F G L L N T F V H I V M Y T Y Y M F S A M G P Q Y Q K Y L
W W K K Y L T T L Q M V Q F I L I M V H A F Q L L F I D C N Y P K
A F V W W I G M H A V M F F F L F N E F Y K A A Y R S R M M K K
N G A L A N G H A K P N G Y C K S I N A H D D L A M P Q T T E A T
A T A T P A S K A N G S S T P P S N G H A N G V E N V Y K Q V A N
G S A H K G S N G G L S N G Y A T K L L D D A S Q E L K Q R K T P
K

CG31523

RNAi trigger

GCCATTTTCACCCACCAGTTCCAGCTTTTGTTCCGCGAATGTGATTACCCCAAGGGCT
TCATGGTCTGGATCGGATTGCACGGCGTTATGTTCCTGTTCCTGTTCTCCGACTTTTA
CAAAGCGAAATATCTCAACGCAGCTCGACGTCGCCGGCAGGCGGTCAAGGCCAATG
GCTATGCCAACGGCTCGGCTTCGAATGGGCACAGCAAGCATCTTGGAGAGGGTGAC
GCCCTGATTGCAAACGGTTGCAATACTGGGGCTTGCATGCCGGTGATGGAGGACGA
GTACGTCAAAAGCAAGGGGCAATCG

Coding Sequence:

ATGGCAATCATACTGCAAGAAGCCCAGAAGTGGTATCGGGACCTAATGGACAATAA
AAGCGATCCTCGCGTTAACGACTTCTTTCTGCTATCCTCGCCACTGCCCACACTTTGT
ATGTGCATATTCTATGCCTACTTCAGCAAATCCCTGGGGCCCAGGCTGATGGCCAAA
CGGAAACCGATGGAGTTGCGATCTGTTTTAGTTGTATACAACGCGATACAAACAATA
TTTAGCGCGTGGATCTTTTATGAGTACTTAATGAGCGGATGGTGGGGCCACTACAGC
TTAAAGTGTCAGCCCGTTGACTATAGCACCACTGGCTTGGCCATGCGGATGGTCAAC
ATTTGCTGGTGGTACTACATATCTAAGTTCACCGAGTTCTTCGACACACTCTTCTTCA
TTCTGCGAAAAAAGAACGAGCACGTCTCCACACTACACGTCATCCATCACGGCTGC
ATGCCGTTCTCCGTCTGGATGGGTCTGAAGTTTGCGCCAGGCGGTCATAGCACTTTC
TTTGCCCTTCTCAACTCGTTTGTGCACATTGTAATGTATTTCTACTACATGATCGCCG
CCATGGGTCCAAAGTACCAGAAGTACATCTGGTGGAAGAAGTACCTGACCACCTTC
CAAATGGTTCAATTCGTGGCCATTTTCACCCACCAGTTCCAGCTTTTGTTCCGCGAAT
GTGATTACCCCAAGGGCTTCATGGTCTGGATCGGATTGCACGGCGTTATGTTCCTGT

TCCTGTTCTCCGACTTTTACAAAGCGAAATATCTCAACGCAGCTCGACGTCGCCGGC
AGGCGGTCAAGGCCAATGGCTATGCCAACGGCTCGGCTTCGAATGGGCACAGCAAG
CATCTTGGAGAGGGTGACGCCCTGATTGCAAACGGTTGCAATACTGGGGCTTGCATG
CCGGTGATGGAGGACGAGTACGTCAAAAGCAAGGGGCAATCGAATGGTGCCTACAA
GGAGGGCTTCTTCAAGGAGGGCGTGTTATCCAACAACGATGCCATCTTCAACCCGGA
CAGCAGTAGTTCTAGTCTGCACCAGCGCAAAGTCAAATAA

Protein Sequence:

M A I I L Q E A Q K W Y R D L M D N K S D P R V N D F F L L S S P L
P T L C M C I F Y A Y F S K S L G P R L M A K R K P M E L R S V L V
V Y N A I Q T I F S A W I F Y E Y L M S G W W G H Y S L K C Q P V
D Y S T T G L A M R M V N I C W W Y Y I S K F T E F F D T L F F I L
R K K N E H V S T L H V I H H G C M P F S V W M G L K F A P G G H
S T F F A L L N S F V H I V M Y F Y Y M I A A M G P K Y Q K Y I W
W K K Y L T T F Q M V Q F V A I F T H Q F Q L L F R E C D Y P K G
F M V W I G L H G V M F L F L F S D F Y K A K Y L N A A R R R R Q
A V K A N G Y A N G S A S N G H S K H L G E G D A L I A N G C N T
G A C M P V M E D E Y V K S K G Q S N G A Y K E G F F K E G V L S
N N D A I F N P D S S S S S S L H Q R K V K •

CG2781

RNAi trigger

CTCGCCGTTCCCCACGATAGCTATCAGTCTGACATACGCGTACATCGTAAAGGTACT
CGGACCCAAGCTCATGGAAAACAGAAAGCCTTTCGAGCTCCGCAAAGTCTTGATCG
TCTACAATGCAGCGCAGGTTATCTTCAGTGCCTGGCTTTTTTACGAGTCCTGCATAGG
TGGATGGCTCAACGGCTATAACTTGCGATGCGAGCCCGTTAATTATTCCTACTCGCC
CAAAGCTATTCGTACCGCTGAGGGCTGCTGGTGGTACTACTTCTCAAAGTTCACCGA
GTTCTTCGACACCTTCTTCTTTGTGATGCGTAAGCGGTACGATCAAGTGTCCACGTTG
CACGTGATCCACCACGGCATCATGCCCGTCTCCGTCTGGTGGGGCGTCAAGT

Coding sequence:

ATGGATTACCTCACAATGTTCTACGACGGCTGGCGGGATCTGATGGACAACAAATCG
GATCCGCGAACACGAGACTACCCGCTCATGAGCTCGCCGTTCCCCACGATAGCTATC
AGTCTGACATACGCGTACATCGTAAAGGTACTCGGACCCAAGCTCATGGAAAACAG
AAAGCCTTTCGAGCTCCGCAAAGTCTTGATCGTCTACAATGCAGCGCAGGTTATCTT
CAGTGCCTGGCTTTTTTACGAGTCCTGCATAGGTGGATGGCTCAACGGCTATAACTT

GCGATGCGAGCCCGTTAATTATTCCTACTCGCCCAAAGCTATTCGTACCGCTGAGGG
CTGCTGGTGGTACTACTTCTCAAAGTTCACCGAGTTCTTCGACACCTTCTTCTTTGTG
ATGCGTAAGCGGTACGATCAAGTGTCCACGTTGCACGTGATCCACCACGGCATCATG
CCCGTCTCCGTCTGGTGGGGCGTCAAGTTCACCCCCGGTGGCCACTCGACCTTCTTT
GGCTTCCTGAACACCTTCGTGCACATCTTCATGTACGCCTACTACATGCTGGCCGCC
ATGGGACCCAAGGTGCAGAAGTACCTGTGGTGGAAGAAGTACCTGACAGTGATGCA
GATGATCCAGTTCGTCCTGGTTATGGTGCACTCCTTCCAGCTGTTCTTCAAGAACGAC
TGCAACTACCCAATTGGCTTCGCCTACTTCATCGGAGCCCATGCGGTTATGTTCTACT
TCCTGTTCTCAAACTTCTACAAGCGTGCCTATGTGAAACGCGATGGCAAGGACAAAG
CGTCGGTGAAGGCAAACGGACATGCCAACGGACACGTGAAGGCGCTCAAGGACGG
CGATGTGGCGCCAACCAGCAACGGCCAGGCCAACGGCTTCCACAACACCTTCTCCA
AGTTCACCACGGATATGTGCAATCCGGCCCTGAACTCCTCGACGAGACAGCGTGTGC
TCGTGAACGCCGGCAACAAGTGA

Protein Sequence:

M D Y L T M F Y D G W R D L M D N K S D P R T R D Y P L M S S P F
P T I A I S L T Y A Y I V K V L G P K L M E N R K P F E L R K V L I V
Y N A A Q V I F S A W L F Y E S C I G G W L N G Y N L R C E P V N
Y S Y S P K A I R T A E G C W W Y Y F S K F T E F F D T F F F V M R
K R Y D Q V S T L H V I H H G I M P V S V W W G V K F T P G G H S
T F F G F L N T F V H I F M Y A Y Y M L A A M G P K V Q K Y L W W
K K Y L T V M Q M I Q F V L V M V H S F Q L F F K N D C N Y P I G
F A Y F I G A H A V M F Y F L F S N F Y K R A Y V K R D G K D K A
S V K A N G H A N G H V K A L K D G D V A P T S N G Q A N G F H N
T F S K F T T D M C N P A L N S S T R Q R V L V N A G N K

CG3971

RNAi trigger

AATCCGCCCAGGTTCATCTCGATGATCATCACGTCGCTGCAGCTGGCTCAGATGATC
ATCGGATGTGCGATCAACGTGTGGGCCAATGGCTTCCTGAAGACGCACGGCACCTC
CTCCTGTCACATCTCGCAGCGCAACATCAATCTGTCGATCGCCATGTACTCCAGCTA
CTTTGTCCTGTTCGCGCGCTTCTTCTACAAGGCGTACCTGGCGCCCGGTGGTCACAA
GAGCCGGCGCATGGCCGCCTCCCTGGCAGCCCAGAATGTGGTCAAGCAGTCGAGCA

Coding sequence:

ATGATCAACATGGACATTAGCGTTACGCCCAATTACTCGTACATCTTCGACTTCGAG
AACGATTTCATACACCAGCGGACGCGCAAATGGATGCTAGAGAACTGGACATGGGT

GTTCTACTACTGCGGCATCTACATGCTGGTCATCTTCGGTGGTCAGCACTTTATGCAA
AATCGGCCTAGATTCCAATTGCGCGGCCCGCTGATCATATGGAACACCTTGCTGGCC
ATGTTCAGCATCATGGGTGCTGCCCGAACGGCGCCGGAGCTGATTCACGTGCTGCGT
CACTACGGACTCTTCCACTCCGTCTGCGTGCCCAGCTACATCGAGCAAGATCGCGTG
TGCGGCTTCTGGACCTGGCTCTTCGTGCTCAGCAAACTGCCGGAGCTGGGCGACACG
ATATTCATTGTGCTCCGCAAGCAGCCACTCATCTTTTTGCACTGGTATCACCACATCA
CCGTGCTGATCTACTCGTGGTTCAGCTATACGGAATATACCAGCTCTGCCCGCTGGT
TCATCGTGATGAACTACTGCGTGCACTCGGTGATGTACAGCTACTATGCTCTGAAGG
CGGCCCGCTTCAATCCGCCCAGGTTCATCTCGATGATCATCACGTCGCTGCAGCTGG
CTCAGATGATCATCGGATGTGCGATCAACGTGTGGGCCAATGGCTTCCTGAAGACGC
ACGGCACCTCCTCCTGTCACATCTCGCAGCGCAACATCAATCTGTCGATCGCCATGT
ACTCCAGCTACTTTGTCCTGTTCGCGCGCTTCTTCTACAAGGCGTACCTGGCGCCCGG
TGGTCACAAGAGCCGGCGCATGGCCGCCTCCCTGGCAGCCCAGAATGTGGTCAAGC
AGTCGAGCAGCCCCCAGCAGGCCAGCGAGAGCTCCAAGTTCATCGGCGCTGGCGAG
GATCAGGCAGCCTATCTGCGCAAGGCCAAGGCGCAGTAA

Protein Sequence:

M I N M D I S V T P N Y S Y I F D F E N D F I H Q R T R K W M L E N
W T W V F Y Y C G I Y M L V I F G G Q H F M Q N R P R F Q L R G P
L I I W N T L L A M F S I M G A A R T A P E L I H V L R H Y G L F H
S V C V P S Y I E Q D R V C G F W T W L F V L S K L P E L G D T I F
I V L R K Q P L I F L H W Y H H I T V L I Y S W F S Y T E Y T S S A R
W F I V M N Y C V H S V M Y S Y Y A L K A A R F N P P R F I S M I I
T S L Q L A Q M I I G C A I N V W A N G F L K T H G T S S C H I S Q
R N I N L S I A M Y S S Y F V L F A R F F Y K A Y L A P G G H K S R
R M A A S L A A Q N V V K Q S S S P Q Q A S E S S K F I G A G E D Q
A A Y L R K A K A Q

CG5278

RNAi trigger

ACGAGGGCGTGGTCATCTGGCAGTACTACAGCTGGCGGTGCCAGCCGGTTGATTGGT
CCCGCACACCCAAGGCGTATCGTGAGGCGCGCGTCGTCTATGTCTACTATCTGGCCA
AGATTACCGAGCTGCTGGACACCATTTTCTTTGTGCTGCGCAAGAATGACCGACAGG
TGACCTTCCTGCATGTCTACCATCACACGGTGATGCCCATGATCAGCTGGGGCACCA
GCAAGTACTATCCCGGCGGACATGGCACCTTCATTGGCTGGATCAACTCCTTCGTGC
ACATTATCATGTACTCCTACTACTTCCTCTCCGCCTTCGG

Coding Sequence:

ATGGCAGCCGTCAACGCAACACAGGTGGACTATTGGAATTTTCTGTTCACCGATCTG
GCAGATCCCCGCACCAATGACTGGTTCCTGATCAAGTCGCCACTGCCCCTGCTGGGC
ATCCTGGCCTTCTATCTCTTCTTCGTGCTGTCCTGGGGACCCAAGTTCATGAAGGATC
GCAAACCCTTCAAGCTGGAACGCACCCTGCTCGTCTACAATTTCTTCCAGGTAGCCC
TCAGTGTTTGGATGGTCTACGAGGGCGTGGTCATCTGGCAGTACTACAGCTGGCGGT
GCCAGCCGGTTGATTGGTCCCGCACACCCAAGGCGTATCGTGAGGCGCGCGTCGTCT
ATGTCTACTATCTGGCCAAGATTACCGAGCTGCTGGACACCATTTTCTTTGTGCTGCG
CAAGAATGACCGACAGGTGACCTTCCTGCATGTCTACCATCACACGGTGATGCCCAT
GATCAGCTGGGGCACCAGCAAGTACTATCCCGGCGGACATGGCACCTTCATTGGCT
GGATCAACTCCTTCGTGCACATTATCATGTACTCCTACTACTTCCTCTCCGCCTTCGG
ACCCCAGATGCAGAAGTATCTGTGGTGGAAGAAGTACATCACCAACCTGCAGATGA
TCCAGTTCTGCTGCGCCTTCATCCACCAGACGCAGCTGTTGTACACGGATTGTGGGT
ATCCCAGGTGGTCGGTGTGCTTTACCCTGCCCAACGCGGTGTTCTTCTACTTCCTGTT
CAACGACTTCTACCAGAAGTCCTACAAGAAGAAGCAGGCTGCTGCCAAGGAGAAGG
CTCTGTCGGCGGATAACAATAATGATGGATGTGCCAAGGACCTCAACAAGGCAATT
CAGCTGCAGCAGGAGAAGCAGAAGGCCTTGTAA

Protein Sequence:

M A A V N A T Q V D Y W N F L F T D L A D P R T N D W F L I K S P
L P L L G I L A F Y L F F V L S W G P K F M K D R K P F K L E R T L
L V Y N F F Q V A L S V W M V Y E G V V I W Q Y Y S W R C Q P V D
W S R T P K A Y R E A R V V Y V Y Y L A K I T E L L D T I F F V L R
K N D R Q V T F L H V Y H H T V M P M I S W G T S K Y Y P G G H G
T F I G W I N S F V H I I M Y S Y Y F L S A F G P Q M Q K Y L W W K
K Y I T N L Q M I Q F C C A F I H Q T Q L L Y T D C G Y P R W S V C
F T L P N A V F F Y F L F N D F Y Q K S Y K K K Q A A A K E K A L
S A D N N N D G C A K D L N K A I Q L Q Q E K Q K A L

GH22993 (full length sequence)

CGATTTCCACACAAAATTCCTAGTCTGCTGCCGAGTACTACCACTGCAAAACGAACT
CAAATTCGAAAACCCGCTTGGCACCGCATCAAATGCACTAAGTCCCAAGCACTAAA
ATTATCCAGATCAGAAGCAAAACCAAGACTTTAACCACCAACCAATTGTACAATCA
AAAATCAGCAAGGAACACGAACTACAAACCAAAGGAATAGTAGGACTCCGCAGTC
GGCGACAAACAGTAACCACTACTCAACAACACAATGGCCTTAATTATGAAATACAT
CGACAGCATAAGCAGATACATGGACTCACATAGTGACTCGAGGACAAAGGGCTGGC
CCATGATGTCGTCCCCATTCCCCACACTGGCCGTTTGCCTCACATATGTCTATCTGGT
CAAGGTGCTGGGACCCCGATTAATGGAGAACCGAAAGCCTCTGAATCTGCAAAACA
CTCTAGTCATGTACAATGCCATACAGGTTGTATTTAGCGCATGGCTGTTTTATGAGAT
AGGCATTTCCGGTTGGCTAACTGGACATTATAGCTTCCGATGTCAGCCAGTCGACTA
TAGTAATAATCCTAGAACGTTAAGGATGGTTCATGCCTGCTGGTGGTACTACTTCTC
GAAGTTTACAGAATTCATGGATACGATTTTCTTCGTGTTGCGCAAGAAGTCGAGCCA
AGTTACCACACTGCACGTCATCCATCACGGCTGCATGCCCATGTCGGTCTGGTTCGG
CGTCAAGTTCACACCGGGTGGTCACAGCACCTTCTTCGGCCTGCTCAACACCTTCGT
CCACATCGTGATGTACACCTACTACATGTTCTCCGCCATGGGTCCGCAGTACCAGAA
ATACCTCTGGTGGAAGAAGTACCTGACCACTCTGCAGATGGTCCAGTTCATCTTGAT
CATGGTGCACGCCTTCCAGCTGCTCTTCATCGATTGCAACTACCCGAAGGCGTTCGT
CTGGTGGATTGGCATGCATGCCGTCATGTTCTTCTTCCTGTTCAACGAGTTCTACAAG
GCAGCCTACAGGAGCCGCATGATGAAAAAGAACGGAGCCCTGGCCAACGGACACG
CCAAGCCCAATGGCTACTGCAAGAGCATCAACGCCCATGACGATCTGGCCATGCCA
CAGACCACGGAGGCGACAGCGACAGCGACGCCCGCATCCAAGGCCAACGGGAGCA
GCACCCCGCCGAGCAATGGACATGCCAATGGGGTGGAGAACGTCTACAAGCAGGTG
GCCAACGGCAGCGCGCACAAGGGGTCCAACGGAGGACTGAGCAACGGGTATGCCA
CCAAGCTGCTGGACGACGCCTCCCAGGAGCTGAAACAGCGGAAGACGCCGAAATAA
TATCGGGGCAAGGAGGGGCGGCGCCAGCGCAGAGCTTAGAACACGGCGAGGGCCA
GTAGCGCCAGGAGCTTCCTGAAAAGGAGCGGAGCTCCAAAAGTCATTTCCATAGGC
ACACACAACTTCATTTAACTCACTACTACTACTTCACATCGCAGACCATCCACTTTCG
AAATTCCAAGCCGGGACCGACGTCCATGGGTAGCTGGGGAAATATGCTATGTAAGC
GCGTGAAAACATTGTTAATCCTGACTGCCGAGGGCATTGCGAGGCGGCTTCTCCCAC
AGACCCAAATCCATACACCTGAACATAGAAATATGCGAATGGTCGCTAGCATACCC
AATGGGTGCACCGCACCCATTGAGGCACATGCAAATGTATAGGCATTGCGAATGCA
TATGCACCCATTCTCTAGCCTACTTTTTAAATACGTTTTGTTGCTTCTTTTTCTTTTTT
AGTTGTACAAAGCTTACTTATCATAAGTATCCCAGCGGATTGTATAATACAAGAACA
AATATCGCAAACCAAGATTTTAACTAATCGAAAACGCTAATGGAAGCCCAGCAAAA
TCAATTTAAAAATAAAACTAATGCCAACGCACTTGGATTGTGTGTGAGCAAAGCATT
CAACTATTTAAGCCTAGAATTTAATTTATATGTGTATATTTGCATACATATATTATTG
AAGTAATTGAAAACAGCAGCAGCAGCCTACGGCAAGTTTTGTAAATGTATAGAGCA
AAAACAAGTGTAATTTAAAATTAGACACAATAATTATTAAAAATAAAAGTCATTCA
AAATAAAAAAAAAAAAAAAAA